Econometrics Assignment 2

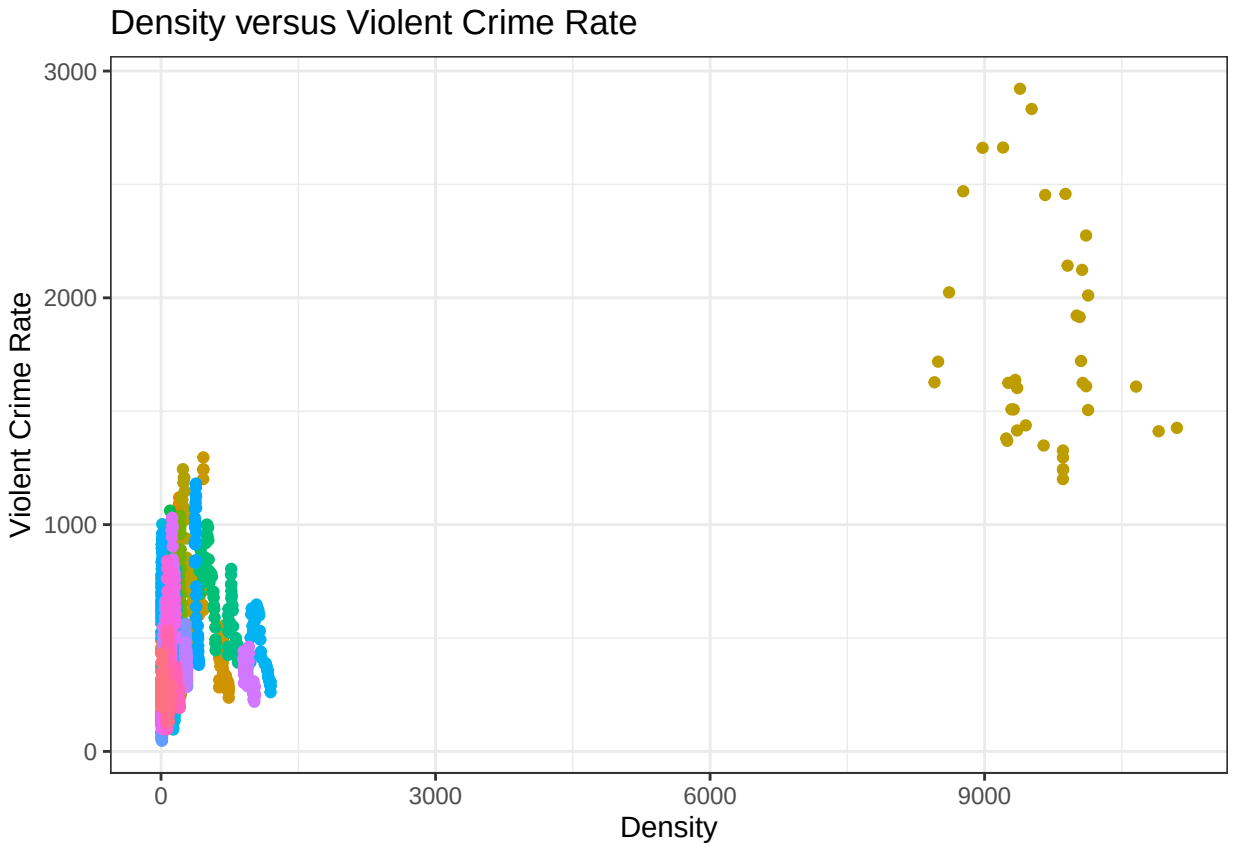
Ruoheng Du

04/12/2023

```
df <- read.csv(file = 'assignment2.csv')
```

Question 1.1

```
ggplot(data = df, mapping = aes(x = density, y = vio,color = state)) +
  geom_point(stat = 'identity', fill = "#4E84C4") +
  ylab('Violent Crime Rate') +
  xlab('Density') +
  ggtitle('Density versus Violent Crime Rate') +
  theme_bw() +
  theme(legend.position = 'none')
```



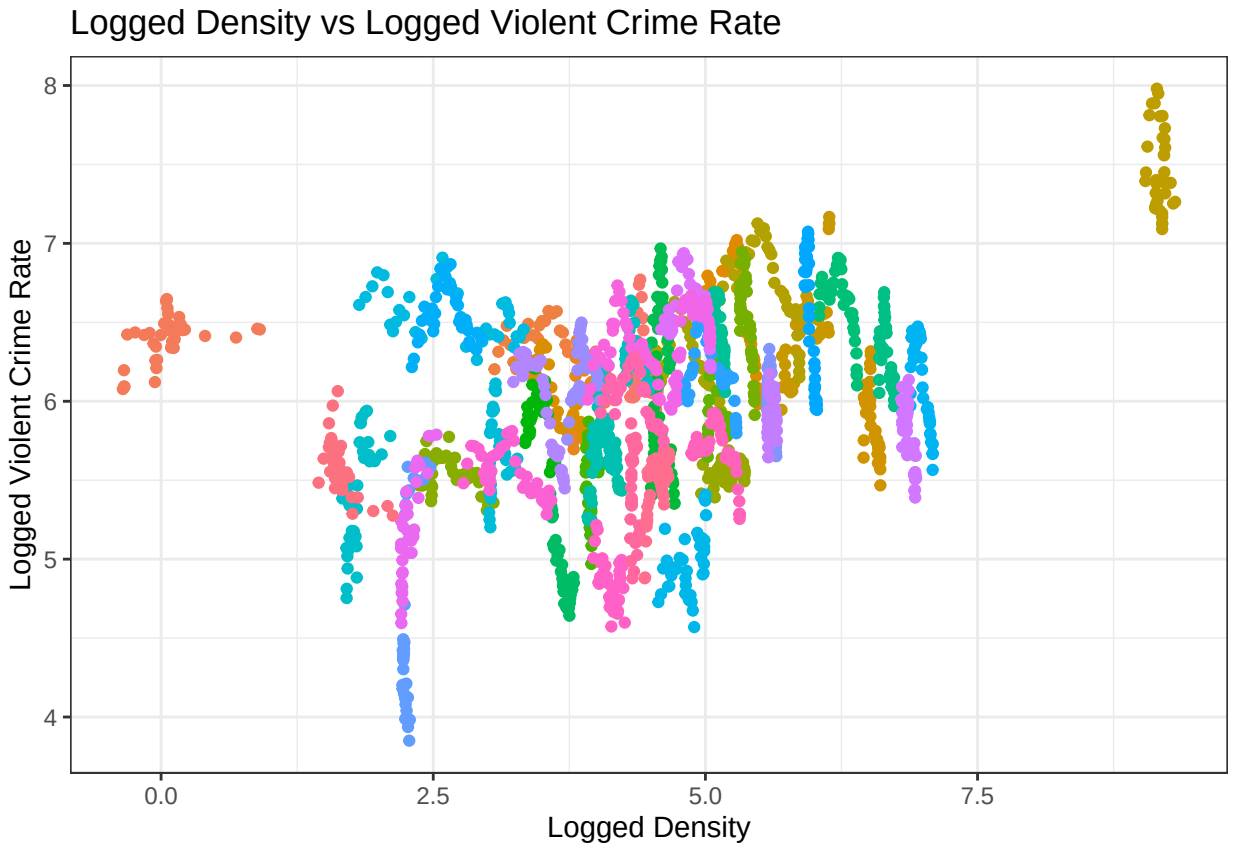
DISCUSSION It is hard to discuss the possible result because there's an outlier state, which make it difficult to observe the correlation between density and violent crime rate. Besides, all the data points are clustered in the lower left side, making it hard to observe the result.

```
df[which.max(df$density),]$state
```

```
## [1] "District of Columbia"
```

DISCUSSION The outlier is the District of Columbia.

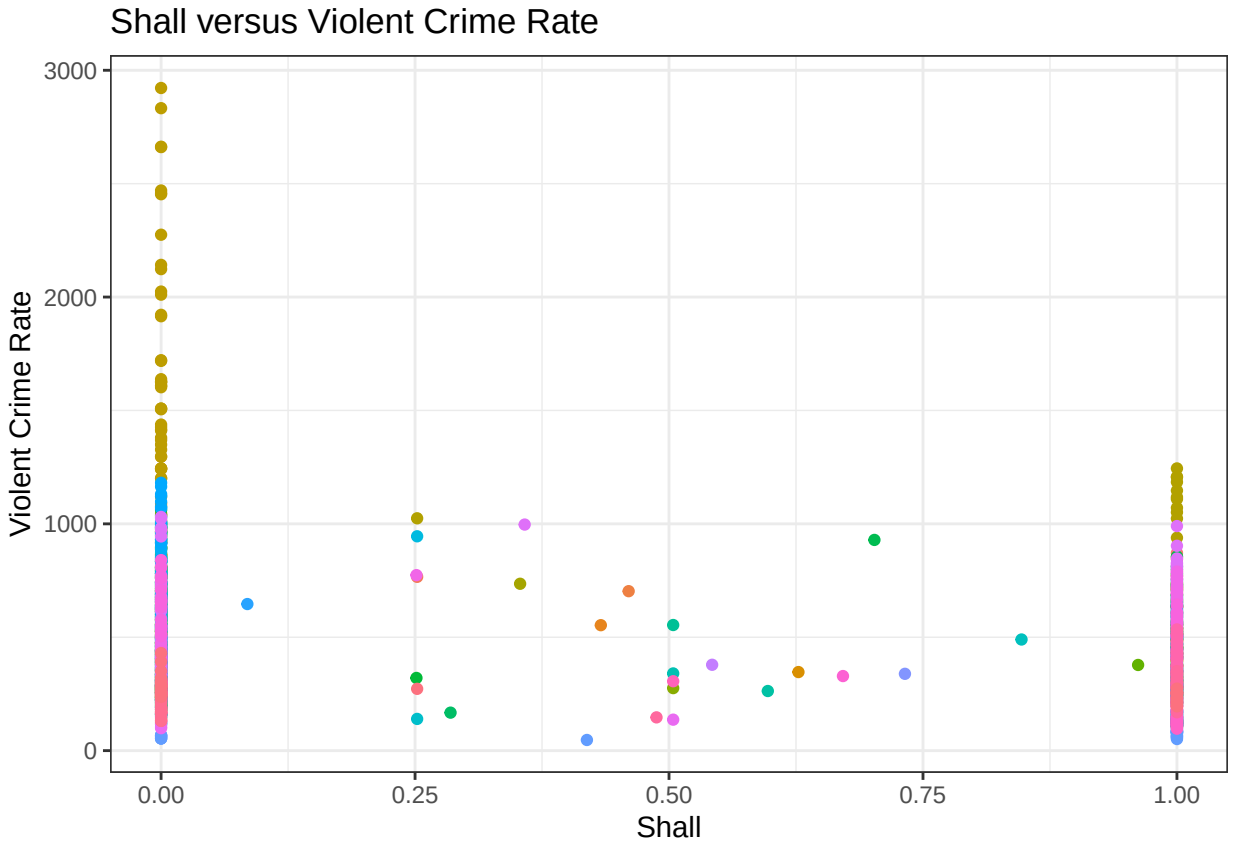
```
ggplot(data = df, mapping = aes(x = log(density), y = log(vio), color = state)) +
  geom_point(stat = 'identity', fill = "#4E84C4")+
  ylab('Logged Violent Crime Rate') +
  xlab('Logged Density') +
  ggtitle('Logged Density vs Logged Violent Crime Rate') +
  theme_bw()+
  theme(legend.position = 'none')
```



DISCUSSION It is useful in transforming the variables as now the scatterplot is more accessible. There is no longer an outlier standing way away from the majority of the points, and it can be inferred that there is a positive relationship between density and violent crime rate.

Question 1.2

```
ggplot(data = df, mapping = aes(x = shall, y = vio, color = state)) +
  geom_point(stat = 'identity',fill = "#4E84C4") +
  ylab('Violent Crime Rate') +
  xlab('Shall') +
  ggtitle('Shall versus Violent Crime Rate') +
  theme_bw()+
  theme(legend.position = 'none')
```



DISCUSSION It is possible that there is a negative relationship between shall and violent crime rate, but due to the influence of the outlier discussed above, it is hard to find a clear correlation.

Question 1.3

```
violent.pd <- pdata.frame(df, index = c("state","year"),
                          drop.index = TRUE, row.names = TRUE)
head(violent.pd, 3)
```

```
##              stateid      vio      mur      rap      aga      bur     auto
## Alabama-1977       1 414.4444 14.20054 25.17615 278.2656 1135.528 280.7317
## Alabama-1978       1 419.0807 13.33511 25.49439 281.1598 1229.316 302.4853
## Alabama-1979       1 413.3192 13.15999 27.51393 263.1467 1287.265 320.0849
##              shall  density    rpcpi    rpcui    rpcim  pbm1019  pbm2029
## Alabama-1977     1 74.55240 9561.056 75.90759 166.6667 2.997776 1.980452
## Alabama-1978     1 75.56670 9914.110 55.21473 157.9755 2.906960 2.025446
## Alabama-1979     1 76.24532 9833.333 57.85124 183.1956 2.817431 2.068932
##               pwm1019  pwm2029
## Alabama-1977 6.832654 6.281442
## Alabama-1978 6.675037 6.299578
## Alabama-1979 6.522353 6.334569
```

```
violent.pd <- mutate(violent.pd, lvio = log(vio))
violent.pool <- plm(lvio ~ shall, data = violent.pd, model = 'pooling')
summary(violent.pool, vcov = vcovHC)
```

```
## Pooling Model
##
```

```
## Note: Coefficient variance-covariance matrix supplied: vcovHC
##
## Call:
## plm(formula = lvio ~ shall, data = violent.pd, model = "pooling")
##
## Balanced Panel: n = 51, T = 38, N = 1938
##
## Residuals:
##      Min.   1st Qu.    Median   3rd Qu.      Max.
## -2.139114 -0.383608  0.038437  0.406258  1.858520
##
## Coefficients:
##              Estimate Std. Error t-value  Pr(>|t|)
## (Intercept)  6.121434   0.084082 72.8028 < 2.2e-16 ***
## shall       -0.314943   0.107348 -2.9338  0.003387 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Total Sum of Squares:    684.78
## Residual Sum of Squares: 637.73
## R-Squared:      0.068716
## Adj. R-Squared: 0.068235
## F-statistic: 8.60746 on 1 and 50 DF, p-value: 0.0050429
```

DISCUSSION On average, to increase shall by 0.1 of the year, violent crime rate is estimated to decrease by 3.149%. This support the previous view of negative relationship that more shall and less violent crime. With the t-statistic of -2.9338 and p-value of 0.003387, the shall coefficient is statistically significant at any reasonable significance level.

Question 1.4

```
violent.fe <- plm(lvio ~ shall, data = violent.pd, model = 'within', effect =
                    'individual')
summary(violent.fe, vcov = vcovHC)
```

```
## Oneway (individual) effect Within Model
##
## Note: Coefficient variance-covariance matrix supplied: vcovHC
##
## Call:
## plm(formula = lvio ~ shall, data = violent.pd, effect = "individual",
##     model = "within")
##
## Balanced Panel: n = 51, T = 38, N = 1938
##
## Residuals:
##       Min.    1st Qu.     Median    3rd Qu.       Max.
## -0.7044334 -0.1435709 -0.0032302  0.1474221  1.0757576
##
## Coefficients:
##        Estimate Std. Error t-value Pr(>|t|)
## shall -0.033673   0.038858 -0.8666   0.3863
##
## Total Sum of Squares:    99.675
```

```
## Residual Sum of Squares: 99.362
## R-Squared:      0.0031396
## Adj. R-Squared: -0.023817
## F-statistic: 0.750942 on 1 and 50 DF, p-value: 0.39032
```

DISCUSSION When adding fixed effects for the different states, on average, a 0.1 increase in shall in the year is associated with a decrease of violent crime rate by 0.33673%. This number is much higher (i.e., smaller effect in absolute sense) compared with the situation that we do not have state fixed effects. However, with the t-statistic of -0.8666 and p-value of 0.3863, the shall coefficient is not statistically significant at any reasonable significance level. Besides, the adjusted r-squared value is negative and the p-value of the F-statistic is above 0.1 (i.e., 10% significance level). So, this model is not good.

Question 1.5

```
violent.fte <- plm(lvio ~ shall, data = violent.pd, model = 'within', effect = 'twoways')
summary(violent.fte, vcov = vcovHC)
```

```
## Twoways effects Within Model
##
## Note: Coefficient variance-covariance matrix supplied: vcovHC
##
## Call:
## plm(formula = lvio ~ shall, data = violent.pd, effect = "twoways",
##     model = "within")
##
## Balanced Panel: n = 51, T = 38, N = 1938
##
## Residuals:
##       Min.    1st Qu.     Median    3rd Qu.       Max.
## -0.7235560 -0.1094477  0.0022864  0.1071006  1.2091013
##
## Coefficients:
##       Estimate Std. Error t-value Pr(>|t|)
## shall 0.028157   0.050734   0.555    0.579
##
## Total Sum of Squares:    76.326
## Residual Sum of Squares: 76.21
## R-Squared:      0.0015253
## Adj. R-Squared: -0.045995
## F-statistic: 0.308012 on 1 and 50 DF, p-value: 0.58138
```

DISCUSSION When we control the different characteristics across states and time, on average, a 0.1 increase in shall is associated with a increase of violent crime rate by 0.28157%. This positive sign is different from the previous two signs, indicating that there could be a positive correlation between shall and violent crime rate. However, with the t-statistic of 0.555 and p-value of 0.579, the shall coefficient is not statistically significant at any reasonable significance level. Besides, the adjusted r-squared value is negative and the p-value of the F-statistic is above 0.1 (i.e., 10% significance level). So, this model is not good.

Question 1.6

```
violent.fe <- plm(lvio ~ shall + rpcpi + rpcui + rpcim + density + pbm1019 + pbm2029 + pwm1019 + pwm202
summary(violent.fe, vcov = vcovHC)
```

```
## Oneway (individual) effect Within Model
##
```

```
## Note: Coefficient variance-covariance matrix supplied: vcovHC
##
## Call:
## plm(formula = lvio ~ shall + rpcpi + rpcui + rpcim + density +
##     pbm1019 + pbm2029 + pwm1019 + pwm2029, data = violent.pd,
##     effect = "individual", model = "within")
##
## Balanced Panel: n = 51, T = 38, N = 1938
##
## Residuals:
##       Min.    1st Qu.     Median    3rd Qu.       Max.
## -0.7703765 -0.1296136 -0.0071264  0.1228994  1.0915365
##
## Coefficients:
##            Estimate  Std. Error t-value  Pr(>|t|)
## shall   -2.2721e-02  4.2245e-02 -0.5378   0.59075
## rpcpi   -2.9098e-05  1.3687e-05 -2.1260   0.03363 *
## rpcui   -1.2790e-03  2.8223e-04 -4.5318 6.216e-06 ***
## rpcim    1.5394e-04  2.8188e-04  0.5461   0.58506
## density -2.5248e-04  2.2570e-04 -1.1186   0.26344
## pbm1019 -8.5509e-02  9.5229e-02 -0.8979   0.36934
## pbm2029  1.0818e-01  9.7412e-02  1.1105   0.26693
## pwm1019 -1.1477e-01  2.6227e-02 -4.3761 1.274e-05 ***
## pwm2029  3.7443e-04  2.2328e-02  0.0168   0.98662
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Total Sum of Squares:    99.675
## Residual Sum of Squares: 80.278
## R-Squared:      0.1946
## Adj. R-Squared: 0.16929
## F-statistic: 15.6361 on 9 and 50 DF, p-value: 9.5518e-12
```

```
violent.fte <- plm(lvio ~ shall + rpcpi + rpcui + rpcim + density + pbm1019 + pbm2029 + pwm1019 + pwm20
summary(violent.fte, vcov = vcovHC)
```

```
## Twoways effects Within Model
##
## Note: Coefficient variance-covariance matrix supplied: vcovHC
##
## Call:
## plm(formula = lvio ~ shall + rpcpi + rpcui + rpcim + density +
##     pbm1019 + pbm2029 + pwm1019 + pwm2029, data = violent.pd,
##     effect = "twoways", model = "within")
##
## Balanced Panel: n = 51, T = 38, N = 1938
##
## Residuals:
##       Min.    1st Qu.     Median    3rd Qu.       Max.
## -0.6476782 -0.1057866  0.0049937  0.1052378  0.9846603
##
## Coefficients:
##            Estimate  Std. Error t-value Pr(>|t|)
## shall   -3.1669e-02  4.8851e-02 -0.6483 0.516883
## rpcpi   -1.6270e-05  1.9777e-05 -0.8227 0.410812
```

```
## rpcui   -1.3795e-03  4.3411e-04 -3.1778 0.001508 **
## rpcim    2.7479e-04  4.7227e-04  0.5818 0.560750
## density -3.5316e-04  1.9396e-04 -1.8208 0.068806 .
## pbm1019  6.2806e-02  1.0697e-01  0.5871 0.557189
## pbm2029  2.1999e-01  1.0658e-01  2.0640 0.039158 *
## pwm1019 -8.5756e-02  4.0139e-02 -2.1365 0.032770 *
## pwm2029  1.0988e-01  4.3081e-02  2.5505 0.010837 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Total Sum of Squares:    76.326
## Residual Sum of Squares: 66.014
## R-Squared:      0.13511
## Adj. R-Squared: 0.090005
## F-statistic: 4.03258 on 9 and 50 DF, p-value: 0.00062891
```

DISCUSSION Beta1 becomes negative again and it changes a lot compared with the original regression, but it is not statistically significant at any reasonable significance level. Compared with the pooled model, the four regression models with state/time fixed effects are more credible because they control for omitted variable bias. Compared with FE and TFE without controls (i.e., in Question 1.4 & 1.5), which have negative adjusted r-squared values, these with controls (i.e., in Question 1.6) have higher adjusted r-squared values of about 17% and 9%. However, the t-statistics of the shall coefficient in these four models with fixed effects indicate that they are all statistically insignificant. Overall, I would say that there is no significant effect of shall-issue laws on the violent crime rate.